PROJECT REPORT

2.6.2.24

Progress Report: _____ Final Report: X Project Number: 82-1032

Small-Scale: _____ Large-Scale: X Funding Year: _____

Transfer Grant: X (Priority: MIA) OISE: _____

New: X Continuing: _____ Terminated: X Re-funded: _____

Project Title: Improving Problem Solving Through Cooperative Learning

Investigator(s): J. A. Ross, F. J. Mavnes, D. Raphael, F. G. Robinson

Budget Unit: FS&R/Trent Valley Centre Date Rec'd in F&SR: Jan 30/87
/Northeastern Centre

ABSTRACT

For large-scale projects, please provide in the space below an abstract of not more than 300 words, single-spaced, written in a language that may be understood by the informed layperson. For small-scale projects, an abstract per se may not be appropriate, and the report itself may be started below.

Cooperative learning group approaches to instruction have been highly successfully in promoting social development goals and in the achievement of low level cognitive objectives such as basic arithmetic skills. There have been very few studies in which cooperative learning has been applied to the more complex cognitive objectives called for in current Ministry Guidelines.

In this project one approach to cooperative learning (Student Teams Achievement Divisions) was applied to one approach to problem solving instruction which has a demonstrated record of effectiveness. The final report of the project contains three sections:

Section 1 reports the results of two quantitative studies. Study 1 compared two teaching methods: a whole class approach and a cooperative learning group method. The results showed that students assigned to both methods substantially outperformed students assigned to the control group. Study 2 was a replication of Study 1 involving a longer period of instructional time (3 months rather than 2 months) and a broader range of problem solving objectives. The results were the same: the two methods were more effective than the control but the differences between the methods were minimal. Four possible explanations of the findings were proposed.

Section 2 reports the results of a qualitative investigation. Study 3 explored one of the four explanations for the findings from Studies 1 and 2. Two cooperative group classrooms were audio and videotaped over a two month period. The patterns of verbal interaction within the cooperative groups were linked to patterns of student achievement. The classroom which developed a less structured approach to cooperative learning outperformed the classroom with a more structured approach.

Section 3 of the report describes the dissemination activities of the project.

ON04771

RETURN TO OISE

TRANSFER GRANT PROJECT OUPUTS

Funding Year: 1986-87 **Project Number:** 81-1032

Project Title: Improving Problem Solving Through Cooperative Learning

Investigator(s): J. A. Ross, F. J. Maynes, D. Raphael & F. G. Robinson

Budget Unit: FS&R/Trent Valley Centre /Northeastern Centre **Date Rec'd in FS&R:** _____

Please list below under two separate headings, Actual Outputs and Anticipated Outputs. Please provide as much detail as possible including present status, e.g., mimeo for distribution, submitted for publication, etc. If you are not providing copies of actual outputs as appendices to your report, please provide one copy for filing in the OISE Library.

ACTUAL OUTPUTS

1. Deserts and Rainforests. Cooperative learning group materials produced by the OISE Trent Valley Centre and the Durham Region Board of Education, 1986.

2. Deserts, Grasslands and Rainforests. Cooperative learning group materials produced by the OISE Trent Valley Centre and the Victoria County Board of Education, 1986.

3. Problem Solving Exercises for the Junior Division Environmental Studies. Instructional materials produced by the OISE Trent Valley Centre and the Victoria County Board of Education, 1987.

4. Ross, J. A. A test of Bloom's two sigma hypothesis. Paper presented at the Canadian Society for the Study of Education, 1987.

ANTICIPATED OUTPUTS

5. Preparation of an article for publication in an academic journal on the outcomes of Study 1 and 2; this would be a revision of item 4.

6. Preparation of an article for publication in an academic journal reporting the outcomes of Study 3.

SECTION 1: QUANTITATIVE INVESTIGATIONS

There is a growing body of evidence that the problem solving performance of elementary and secondary students can be improved substantially through the use of direct instructional techniques in whole class settings. One set of instructional procedures (elaborated in Robinson, Ross & White, 1985) involves decomposing complex tasks into component operations, making the cognitive processes underlying these operations explicit to students, incrementally modifying students' cognitive structures in a controlled sequence that reduces demands on learners, and consolidating improvements in performance with increasingly demanding practice.

Figure 1 contains an example of a problem solving task (a decision making problem) and a hierarchy of student responses to it--in this instance the responses focus on the framework developed by students for handling information for the problem. Instruction (elaborated in Maynes & Ross, 1984 and Ross & Maynes, 1982) for moving students from one level to another in this scheme consists of having students individually perform the target skill in a real life or curriculum specific context, displaying to all students the lowest level of performance existing in the class, demonstrating in whole class discussion the inadequacies of the existing performance, and stimulating students to collectively develop a more sophisticated response. Practice consists of eliciting or providing rules for performing the skill and consolidating these rules with individual seat work using topics appropriate to the grade and subject in which the instruction occurs.

Figure 1 About Here

Evidence of the effects of the technique have been reported in a variety of grades and subjects. A meta-analysis found a mean effect size of .56 (based on 39 findings from 9 studies involving 2000 students) in favour of this method over incidental teaching of problem solving (Ross & Maynes, 1983). Retention studies (conducted after one month, four months and six months) indicate that the effects of the instruction on problem solving performance tend to endure (Ross & Maynes, 1985).

In this approach within-class variability in student performance on entry to instruction has been problematic for teachers. Solutions attempted include the development of self instructional kits for high ability students, the use of pedagogical asides that amplify instruction for slower students and differentiated practice activities. Robinson, Ross & White (1985) have proposed that grouping strategies, particularly homogenous grouping, might meet student needs more effectively than a whole class method. But there is an emerging literature suggesting that the use of homogenous groups within classes, like ability tracking between classes, has only modest benefits for high ability students and these are vastly outweighed by the serious negative effects on students assigned to low ability groups (Good & Marshall, 1984; Eder & Felmlee, 1984). These data suggest that heterogenous ability groups are preferable.

There is evidence that the manipulation of classroom reward structures is central to the effects of mixed ability groups on achievement (Johnson, et al., 1981; Michaels, 1977; Sharan, 1980; Slavin, 1977, 1983). A meta-analysis of 122 studies found that students in cooperative learning situations consistently outperform students in competitive and individualistic situations; the effect size was .78 (Johnson, et al., 1981).

In the two studies described below, the two innovations (direct instruction of problem solving and cooperative learning groups) were combined. Following Bloom (1984), it was hypothesized that students in the combined method would outperform students in the single innovation, and that the effects of the two methods on achievement would be additive.

STUDY 1

Sample

Subjects were 344 grade 4 social studies students from a large school system in central Ontario; 17 intact classes were randomly assigned to three treatment conditions.

Treatments

1. The whole class method was an adaptation of the direct instruction strategy for teaching problem solving skills described above. The target skills concerned comparison and decision making. Teachers were provided with highly detailed instructional materials which guided classroom discussion and which contained student worksheets and weekly quizzes. Figure 2 provides an example of a student worksheet; the correct answer and suggested feedback to students is given in Figure 3. The typical lesson consisted of 20 minutes of teacher directed dialogue with the whole class, followed for 20 minutes by students working independently on practice worksheets. The reward structure was individualistic: student grades were based on each child's performance on quizzes given approximately weekly. There were two five week units. Teachers attended a half day inservice to familiarize themselves with the materials and teaching techniques.

Figures 2 and 3 About Here

2. The cooperative group treatment followed the procedures for the Student Teams-Achievement Divisions method described in Slavin (1980). The instructional materials were identical to those used in the whole class method. The typical lesson consisted of 20 minutes of teacher directed dialogue with the whole class, as above, followed for 20 minutes by students working on practice worksheets in heterogenous groups consisting of one high ability student, two medium abilities and one low ability student. Students were told to use the worksheet to prepare themselves for the weekly quizzes which each student completed individually. The reward structure was cooperative: students were given incentives meaningful to grade 4 students (e.g., bulletin board recognition) based on the aggregate of group performance, adjusted by ability; (see Slavin, 1980 for the exact procedure). Teachers attended a half day workshop to familiarize themselves with the materials and teaching techniques of problem solving instruction. A second

half day workshop was held on the techniques of cooperative learning.

3. The control group treatment consisted of knowledge driven units in which there were opportunities for teaching problem solving skills but no explicit direction or encouragement was given for doing so.

Instruments

Three versions of each of the problem solving and knowledge instruments were developed; there was a single version of the two affect tests. Alternate versions of the instruments were combined into six test orders which were randomly assigned to students. Tests were administered before, immediately after and one month after the units were taught.

Problem solving skills were assessed using open ended instruments developed in previous investigations. These items were transfer tests in that they involved familiar content not used in the instruction. The instructions for the comparative item were of the form "Compare two sports"; instructions for the decision making item were of the form "Suppose your class is going on a school trip. How should the class decide where to go". Figure 3 contains the coding scheme used to interpret student responses. The pre and post tests were marked by the field test teachers, following a training session that brought the teachers to an average interrater agreement between teachers of .88. Samples of responses were drawn from each class set of tests and independently assessed by trained markers; if the interrater agreement between teacher and trained marker was below 80% the entire class was recoded. All the retention tests were coded by trained markers; the average interrater agreement between markers was .97

Figure 3 About Here

Knowledge was assessed using short answer items that tested important propositions contained in the units. For example "Nomads are people who (a) live in cities, OR *(b) move around from place to place, OR (c) eat nothing but vegetables. Each knowledge instrument consisted of 5 items; Hoyt reliability of the knowledge tests averaged .84.

There were two affect instruments: 24 items measuring intrinsic motivation (Harter, 1981) and 8 items measuring feelings of competence (derived from Bandura, 1977 and Schunk, 1981). Hoyt reliabilities were .91 for the motivation instrument and .70 for the feelings of competence measure.

Analysis

In the design students were nested within classes which were nested in treatments and crossed with sex--sex differences in responding to cooperative situations have been reported by Lindow, Wilkinson & Peterson (1985), Peterson & Fennema (1985) and Webb (1984, 1985). Initially the unit of analysis was the class: a 3 X 2 analysis of covariance, with pretest scores as covariate, was used to analyze posttest and retention test differences between the three treatments and between males and females. A subsequent analysis in which the unit of analysis was the student used a 3 X 7 X 2 analysis of covariance.

Results

The means and standard deviations for the three treatments on the three test occasions are shown for all the instruments in Table 1 with the class as unit of analysis.

Table 1 About Here

When the class was used as the unit of analysis there were no significant posttest or retention test differences on the knowledge and affect instruments.

Four of the six problem solving indicators showed a treatment effect on the posttest: the use of categories in a comparison, $F(2,27)=10.95, p<.01$, generating decision alternatives, $F(2,27)=8.51, p<.01$, developing decision criteria, $F(2,27)=10.22, p<.01$, and summarizing information in a decision problem, $F(2,27)=30.20, p<.01$. In all of these comparisons the whole class and cooperative methods substantially outperformed the control condition. On all the problem solving measures there was a slight, nonsignificant advantage for the whole class over the cooperative treatment. There were no significant sex or sex-treatment interactions. The same pattern of results was found on the retention test scores.

A more complex picture emerged when student was used as the unit of analysis. The means and standard deviations for all three test occasions for all instruments are shown in table 2. For the posttests, there was a treatment effect, in the same direction as in the class analysis, on each of the four problem solving indicators previously discussed. On all the problem solving indicators there was a main effect for teacher nested within treatment: some teachers were consistently more successful in promoting student growth than their peers. On two of these items, using categories in a comparison and generating decision alternatives, there were sex differences: $F(2,308)=5.31, p<.02$ and $F(2,305)=4.93, p<.03$. These differences indicated that males learned more over the course of the treatments than females. The male advantage was particularly pronounced in the cooperative treatment on one indicator; there was an interaction of sex with treatment on using categories in a comparison, $F(2,308)=5.95, p<.02$.

Table 2 About Here

The same pattern of findings favouring the whole class and cooperative treatments over the control held in the retention score data. There were also a number of interactions. These were of minimal interest. For example, scores on interpreting comparative data improved on the retention test, $F(1,308)=15.63, p<.01$; this improvement was largely attributable to a few teachers in the control and cooperative conditions, $F(14,308)=3.72, p<.01$, and males did particularly well in the control condition $F(2,308)=5.10, p<.02$. Other findings showed teacher effects (in using comparative categories and developing decison criteria), and sex by teacher effects (in generating decision alternatives and summarizing decision making information). Of greater interest was the finding that on this last indicator, scores improved from post test to retention test for the whole class treatment but declined for the cooperative treatment.

There were few posttest and retention test differences in the knowledge instruments when the unit of analysis was the student; for example, there was a teacher effect and a sex by treatment interaction on the deserts posttest. There were minor teacher effects on the affect instruments.

In summary the findings from the first study were contrary to Bloom's hypothesis. The cooperative and whole class each outperformed the control condition but the combined innovation (cooperative treatment) was no better than the single innovation (whole class treatment)--in some respects the single innovation was superior.

STUDY 2

Sample

Subjects were 258 grade 4 social studies students from a small school system in central Ontario; 14 intact classes were randomly assigned to the three treatment conditions.

Treatments

The treatments were identical to those of Study 1 except that there were three units rather than two and the duration of the treatments was 15 weeks. Approximately 15% of the instructional materials used in Study 2 were also used in Study 1.

Instruments

Problem solving skills were assessed using open ended instruments in the same format as those used in Study 1. In addition to comparative thinking which was also tested in Study 1, Study 2 measured performance on experimental thinking. Additional items were prepared to assess students' ability to focus comparative problems, write comparative paragraphs, and focus experimental problems. The assessment schemes for these items are given in Figure 5. The knowledge items were of the same form as those used in Study 1 and the affect instruments were identical. Instrument reliability was established using the same procedures developed for Study 1. Interrater reliability of the skill items averaged .95 for the trained markers; Hoyt reliabilties averaged .61 for the knowledge items and averaged .81 for the affect instruments.

Figure 5 About Here

Analysis

The same design was used in Study 2 as in Study 1: a 3 X 2 analysis of covariance when the class was used as unit of analysis, followed by a 3 X 5 X 2 analysis of covariance with the student as unit of analysis.

Results

The results were very similar to those of Study 1. The means and standard deviations for the three treatments on the three test occasions are shown for all the instruments in Table 3 with the class as the unit of analysis.

Table 3 About Here

With one exception, (the control group scored lower than the other treatments on grasslands knowledge), there were no posttest differences on the knowledge and affect instruments.

There were signficant differences on six of the nine items measuring problem solving performance: using categories in a comparison, $F(2,19)=20.04, p<.01$, interpreting a comparison, $F(2,19)=4.60, p<.03$, controlling variables in an experiment, $F(2,19)=47.70, p<.01$, focussing a comparative problem, $F(2,19)=6.42, p<.01$, writing a comparative paragraph, $F(2,17)=19.70, p<.01$ and focussing an experimental problem, $F(2,19)=24.02, p<.01$. In every case the whole class and cooperative methods substantially outperformed the control treatment. There was a slight trend favouring the whole class over the cooperative method. There was one posttest result, on comparative paragraph writing, in which females outperformed males. The same pattern of results was visible on the retention tests: the whole class and cooperative treatments outperformed the control. On the retention tests there was one indicator in which there was a differential decline in the cooperative treatment (comparative paragraph writing) and one item in which there was a differential decline in the whole class method (focussing an experimental problem).

The results when the student was the unit of analysis are displayed in Table 4. The posttest scores on the problem solving items showed that the whole class and cooperative treatments outperformed the control condition. The differences were statistically significant on five of the nine items: using categories in a comparison, $F(2,231)=11.76, p<.01$, controlling variables in an experiment, $F(2,230)=44.66, p<.01$, focussing a comparative problem, $F(2,233)=7.04, p<.02$, writing a comparative paragraph, $F(2,215)=7.58, p<.02$, and focussing an experimental problem, $F(2,234)=14.29, p<.01$. The retention scores on the problem solving items were the same except there was a increase for the whole class on one indicator (using categories in a comparison) and a decline for the same treatment on another (interpreting a comparison). Females did better on several indicators but there were no sex-treatment interactions. There was a consistent teacher nested within treatment effect on virtually all items. The posttest and retention scores on the knowledge and skill items were very similar to the findings from the class analysis.

Table 4 About Here

DISCUSSION

Bloom's hypothesis, that combined innovations would have a more positive impact on student achievement than single innovations, was not confirmed. In both studies the cooperative treatment, which combined direct instruction of problem solving with the use of heterogeneous grouping, performed no better, and in some respects more poorly, than the direct instruction method alone.

One possible explanation for this unexpected finding follows from the counter-hypothesis that combining innovations will overload teachers and lead to incomplete implementation. There was evidence of significant deviation from the intended treatments in both studies. For example, in one cooperative classroom the teacher failed to provide group rewards. But partial or nonimplementation also occurred in the whole class method; for example, one teacher transformed a skill driven unit into a knowledge focussed unit. Observations by the investigators and interviews conducted with teachers at the end of the project suggest that incomplete implementation was the norm for all treatments. Differences in degree of implementation help explain why there was a strong teacher effect in the analyses of covariance but is insufficient as an explanation of the main findings. There may have been a slightly greater tendency for the single innovation to be implemented more completely than the combined innovation but the difference was not so large as to be detected by the investigators.

A more likely explanation is that there is an interaction between type of objective and the effects of cooperative grouping. One view of this interaction suggests that cooperative treatments are more effective for higher order objectives (Johnson, Maruyama, Johnson & Nelson 1981); evidence in partial support of this position is provided by Stodolsky (1984). But a series of other studies suggest that the opposite might be the case, that cooperative approaches are less effective for higher order skills (Bargh & Schul, 1980; Cohen, Kulik & Kulik, 1982; Foster & Penick, 1985; Okabukola & Ogunniyi, 1984; Rich, Amir, Ben-Ari & Merarech, 1985; Slavin & Karweit, 1985). These studies suggest that combining innovations in the present investigation did not have an additive effect on achievement because of a fundamental incompatibility between the innovations.

The present investigation can shed some light on the source of this incompatibility help explain why STAD, the cooperative learning groups treatment used in Study 1 and 2, makes little contribution to learning of problem solving skills.

1. The amount of independent practice may have been insufficient to achieve mastery. When STAD is used for teaching computational skills, students are typically given 20-30 short items to work through in each group session. In contrast the practice exercises in our studies typically involved one or two long problem solving items in each session; (see for example Figure 2). Although it is possible to have shorter and fewer items, the task to be practiced inevitably takes more time, particularly when an attempt is made to integrate problem solving development with acquisition of subject specific knowledge. Burns & Lash (1986) discovered a similar tendency: they found that when math teachers shifted from computational skills to problem solving the

number of items given for practice declined dramatically.

2. The distribution of cognitive demands across the task may have had unintended effects. In Study 1 and 2 cognitive activity was concentrated in the intial phase of the practice task. In Figure 2, for example, the most difficult mental action is undertaken at the outset with the construction of a framework containing alternatives as one dimension with advantages and disadvantages as the other. This massing of demands may have been counterproductive since previous researchers (e.g., O'Donnell, et al., 1985) have found that spreading cognitive activity evenly across the exercise is more effective. The concentration of demands at the beginning may also have increased the consequences of social loafing. Latane, Williams & Harkins (1979) found that in group situations some students will allow others to do most of the work. If students prone to loafing were slower to get going on the exercise, a reasonable hypothesis, the most difficult part of the task would be completed by others, with the result that the loafers would receive practice only on the less demanding problem solving operations.

3. Student learning in groups may have been impeded by the lack of competent/confident students able to give direction to others. When computational skills are the object of instruction it is reasonable to anticipate that in each heterogeneous group there will be at least one person who knows how to do the task on entry. Furthermore many arithmetic skills can be easily mastered by bright students in short periods of time. In cooperative settings these masters become available to help others. One might speculate that the presence of at least one master in the group is an essential condition; for example, Webb (1980) found that math groups which lacked a master experienced motivational problems and tended to be unproductive. In Study 1 and 2 most students performed very poorly on all the skill pretests, and the tasks to be learned were substantially more challenging than the acquisition of computational skills. It is likely that the most competent person in each group was at best an incomplete master, reluctant and unable to provide effective tutoring to others.

4. Helping behaviours may not have occurred in the cooperative groups to the degree required to achieve mastery. Research into unstructured cooperative settings have found that certain types of student-student interactions are critical: providing explanations to others contributes to achivement because it involves rehearsal and elaboration of the material to be learned (Peterson, et al., 1981; Peterson & Swing, 1984; Webb, 1982; Webb & Culligan, 1983); receiving explanations in response to a request also improves performance (Webb, 1982; Webb & Kinderski, 1984), as does providing procedural information on group progress and other metacognitive acts (Peterson, et al., 1984). These effective behaviours are relatively infrequent in groups composed of young children: Peterson & Swing (1984) found that only a few mature and bright students were able to provide good explanations, even on tasks related to lower order objectives.

SECTION 2: QUALITATIVE INVESTIGATIONS

One of the explanations for the unexpected findings in studies 1 and 2 concerned the verbal interactions of students. Previous investigations of cooperative learning groups based on Slavin's techniques (such as STAD--the method used in Study 1 and 2) have not examined the specific interactions of students within the groups.

In contrast, a number of studies have been conducted in which the social interactions in naturally occurring groups have been related to achievement outcomes. These studies tend to have weak instructional designs--if groups are established students are typically told to "work together" and check their responses to low level cognitive tasks with other students before asking for assistance from the teacher. The usual result of such instruction is that there are low levels of student-student interaction and no overall achievement differences between small group and whole class settings (Peterson & Janicki, 1969; Peterson, et al, 1981; Webb & Kenderski, 1984). Despite the absence of main effects these process-product investigations provide clues as to why well designed cooperative groups might facilitate learning and help to identify specific patterns of peer interaction that affect academic growth:

(a) The achievement of high ability students increases if they engage in social interactions in which they give academic explanations to other students (Peterson, et al, 1981; Peterson & Swing, 1984; Swing & Peterson, 1982; Webb, 1982; Webb & Cullian, 1983); formulating an explanation requires a cognitive restructuring that contributes to the learning of the one giving the explanation. Low ability students who give academic explanations benefit in the same way but there are many fewer instances of this type. High ability students who provide procedural information on group progress outperform high ability students who do not (Peterson, et al, 1984): these procedural comments serve a self-pacing and monitoring function that provides metacognitive guidance to the individual making them (usually high ability students only).

(b) The identification of interactions that promote the achievement of lower ability students is much less clear. Some studies (Peterson, et al, 1981; Peterson, et al,); found that receiving academic explanations from a peer is unrelated to achievement. In contrast Peterson & Swing (1984) found a weak relationship between receiving an explanation and academic improvement. In clarification, Webb (1982) discovered that receiving an explanation improved student performance only if an explanation was requested and involved more than a statement of the right answer. Similarly, not receiving an explanation in response to a request for help substantially reduces achievement of lower ability students (Webb & Cullian, 1983; Webb & Kenderski, 1984).

These investigators attribute such findings to the difficulty students have in providing good academic explanations--only mature very bright students are capable of doing so in these studies (Peterson & Swing, 1984). The failure

of peer tutoring in the interactions observed helps explain why the grouping strategies used in this body of research had no advantage over whole class arrangements. This failure can also be attributed to the weakness of the instructional design: there is evidence that peer tutoring is successful only when, in contrast with these projects, there is a detailed analysis of the subject matter to be taught, limited objectives and well structured programs (Cohen, Kulik & Kulik, 1982; Ellson, 1976).

(c) Low ability students who uttered vague requests for help from their peers performed more poorly than lower ability students who ignored the group and worked independently (Peterson, et al, 1984; Webb & Kenderski, 1984). Seeking help was highly correlated with off-task behaviour (also negatively related to achievement). The finding suggests that students may have difficulty articulating requests for assistance in order to access peer support successfully. There is evidence from sociolinguistic studies that children who are effective speakers (i.e., they formulate requests which are sincere, direct, on-task and, if necessary, revised) tend to be higher achievers (Wilkinson & Spinelli, 1983). This correlational finding provides little information about causal direction and no training studies have been conducted designed to improve achievement by upgrading student requests for help.

Study 3 attempted to shed some light on the effects of cooperative learning by intensively examing the interactions of students within the small groups.

Sample

Subjects were 51 grade 4 students in two intact classrooms in a large school system. The classes were taught by "Mary" (an experienced teacher who had been part of the collaborative team that produced the materials developed for Study 1) and "Fred" (an experienced teacher who had been one of the teachers in the cooperative learning groups sample in Study 1 who obtained above average student achievement results). (Both names are pseudonyms.)

Treatments

Both classes used the cooperative method known as STAD (Student Teams-Achievement Divisions); as in Study 1, the specific procedures were from the second edition (Slavin, 1980) rather than the third (Slavin, 1986). (The major difference in the editions is that in the later version the scoring of group work is simplified and competition between teams is de-emphasized.) There was a small modification from Study 1 of the methods used to create the cooperative groups: Almost all of the cooperative literature assumes that it is the high ability child who will be most helpful in giving explanations. Studies of student interaction in natural cooperation settings suggest that a small number of children are responsible for initiating and receiving communications and that ability is only one factor in the makeup of good communicators (Cooper, Marquis & Ayers-Lopez, 1982). Consequently the teachers in Study 3 adjusted the group assignments to ensure that each group had at least one good communicator.

The instructional materials used in Study 3 were very similar to those used in Study 1 except that the worksheets were simplified (in response to teacher feedback from Study 1) and the number of worksheets was increased (in an attempt to increase the number of practice opportunities). These changes increased the length of the units; unlike the teachers of Study 1, neither Study 3 teacher finished the second unit. The duration of the instruction was two months.

There were planned differences in the implementation of the cooperative method. Fred used a highly structured approach. Each student was assigned a specific role in the group process. The roles were: handing out the worksheets, making sure everyone has a turn to speak, handing out the worksheet answers and checking that everyone has done the work right. These roles were rotated among the members of each group after every exercise. The roles were described to students at length and students were reminded throughout the project to attend to their role responsibilities. Previous investigators have found that assigning roles within groups contributes to learning (Johnson, Johnson & Stonne, 1986; McDonald, et. al., 1979; Mesch, Johnson & Johnson, 1986).

Fred also gave much greater attention to the generic task of the worksheets. In the general procedure that was common to all the worksheets students were asked to: (a) read the worksheet and work on it individually for 5-15 minutes; (the purpose of this step was to reduce "social loafing"); (b) share answers with other group members; (c) compare answers to the official answers and (d) discuss any discrepancies between student answers and the official answers. This procedure was prominently displayed and student attention reguarly directed toward it as part of a conscious training program. Devin-Sheehan, Feldman & Allen (1976), in a review of the peer tutoring literature, found some evidence that training tutors contributes to student learning.

Mary was much less structured in her approach to the group activities. She did not assign roles. She gave little attention to the generic procedure used to complete the worksheet; in particular she did not have students engage in 5-15 minutes of independent work at the beginning of each task. In the first few exercises she told students to have one group member record the outcomes of the discussion such that each group would produce a single answer. In practice students filled out their own worksheets so she encouraged students to write down their own answers as the work of the group progressed without necessarily coming to agreement.

The specific treatments were developed in collaboration with the two teachers. Several suggestions made by the investigators were declined. These were: (1) provide additional practice in cooperating with other children through the school day; (cooperative games developed for physical education classes and some of Slavin's arithmetic activities were made available); (2) de-emphasize the competition between groups by awarding bonus points to each group based on class performance (as recommended by Johnson, Johnson & Stonne, 1986), and (3) reward group processes by giving bonus points to groups displaying specific cooperative behaviours such as helping others (as recommended by Mesch, Johnson & Johnson, 1986). The teachers did not object to any of these suggestions--they simply felt that their task was sufficiently

demanding without adding more things to worry about.

Instruments

The knowledge and intellectual skill instruments were identical to those used in Study 1 and were marked by the principal investigators using the same procedures as in the previous study. The affective instruments used in study 1 were deleted and replaced with three items that required students to list 6 people the child would like to have to (a) play games with, (b) be in the same small group with, and (c) get help from. A new question was added to the demographic portion of the instrument to identify children with younger brothers and sisters. Bargh & Schul (1980) found that children with younger siblings were more likely to explain things to their peers. The knowledge and skill instruments were administered before, after and one month after the units were taught. The affective items were administered on the pretest and on the retention test.

Additional achievement data was also available from the unit quizzes administered approximately every two weeks. There were 4 quizzes for the first unit and one quiz from the second.

Data produced through the achievement measures was supplemented with observation of small group interactions [note 1]. Each student in the group to be recorded was provided with an Astatic Model 840 Lavalier microphone. The 4 or 5 (depending on group size) student microphones were attached to a Tascam 106 Audio Mixer. The student microphones were fed into channel 1 of the mixer; channel 2 received the output from a microphone attached to an observer. Students wore numbered stickers on their shoulders. As each student spoke into channel 1, the observer gave the student's number into channel 2. The outputs from the mixer were fed into a Nikko Stereo Tape Deck which recorded output onto Maxell UR 120 tape cassettes. The input was monitored by the observer who wore stereo (Sennheiser HD 414X) headphones during the taping. The observer was able to raise or lower the recording levels to accommodate particularly soft spoken children. An additional output from channel 1 of the mixer was fed into a videocamera and onto video tapes (Scotch Color Plus High Grades T120). The class sessions were approximately 60 minutes in length. Mary's class was recorded 10:30 - 11:30 a.m. Fred's was recorded for the same amount of time but it was split by recess (1:15 - 2:30 p.m.). All recordings were made in the regular classroom; groups not being recorded on a given day engaged in the same cooperative worksheets in the same room.

Each group was recorded twice in the first unit. Students were then assigned to new groups (as recommended in Slavin, 1980) and each of the new groups was recorded once. Cooperative learning group activities were not held in every class session because a portion of class time was spent introducing new material and engaging in other whole class activities.

A Sony model transcriber was modified to allow the observer to focus upon either channel. The student conversations were first transcribed from channel 1 and then the student identifications from channel 2.

RESULTS

Data were collected from 36 observations of Fred's and Mary's classes. Each class was observed for 12 Desert sessions and 6 Rainforests. Since each class had six groups of children, each group was observed twice for Deserts and one each for Rainforests. All groups consisted of four students except for two in Fred's class and one in Mary's class which contained five.

Observation of Implementation

Initially Fred's more structured approach seemed more effective. Students' seemed more task oriented, the room was strikingly more quiet than Mary's and the verbal interaction seemed more orderly and better structured. This perception of the relatively "higher" nature of verbal interaction was corroborated by perusal of the written transcripts of the verbal interaction during the sessions.

However, as the units progressed Fred's class began to show a marked increase in boredom and non-interest in the tasks while Mary's class maintained a high level of excitement, interest in, and enthusiasm for the activities. It also appeared that Fred's class became increasingly less likely to examine the answer sheets, and spent longer portions of their time working silently without interacting with other members of the group. At the urging of Fred the Rainforest unit was terminated a week before the Spring break as he perceived the children were beginning to find the sessions "painful".

Problem Solving Results

Each of the skill instruments consisted of three measures. Analysis of the internal consistencies of the measures indicated that reliabilities were high enough to allow the computing of an overall Comparison and Decision Making score for each child for pretest, posttest, and retention. Table 5 presents scores for these measures for children in Fred's and Mary's classes. While 51 students were enrolled in the two classes, 47 contributed data at all three collections.

Table 5 About Here

A multivariate analysis of variance with repeated measures (pretest, posttest, and retention) revealed significant effects for both Comparison-Making: Teacher [$F(1,45)=4.76$, $p<.05$]; Observation [$F(2,43)=6.71$, $p<.01$] and Teacher by Observation interaction [$F(2,43)=34.01$, $p<.001$] and Decision-Making: Teacher [$F(1,45=80.62$, $p<.001$]; Observation [$F(2,44)=28.84$, $p<.001$] and Teacher by Observation interaction [$F(2,44)=34.01$, $p<.001$]

Univariate tests indicated that Fred's class outscored Mary's on Comparison pretest ($p<.05$) but that Mary's class outscored Fred's on post and retention. For the Decision Making measures, no difference was seen for pretest and Mary's class outscored Fred's on posttest and retention ($p<.05$).

However, since pretest differered significantly between classes, these scores were used as a covariate. Results of this analysis indicated that the

regression of post and retention scores upon pretest (that is controlling for pretest scores) did not weaken the post and retention advantages for Mary's class for both posttest and retention.

A final multivariate analysis examined the effects of sex on pre, post and retention Comparison-Making and Decision-Making scores. None of the t-values associated with the regression of sex upon achievement were significant.

The main achievement results were unambiguous. The low structure method dramatically outperformed the high structured method. Fred's class showed very little if any improvement over the two months. There was a significant decline from pre to post on the comparison measure and no change on the decision making instrument. The retention scores indicated modest improvement from the posttest results, indicating that the treatment had a small delayed impact. In contrast, Mary's students improved substantially from pre to post on both measures and these tests were retained on the final test administration.

Quiz Score Results

Students took three quizzes for the Deserts units and one for Rainforests. These quizzes served primarily as formative evaluation instruments allowing Fred and Mary to monitor student learning. The quiz scores combined knowledge and skill components in a single score. Table 6 presents mean quiz scores for Fred and Mary's classes.

Table 6 About Here

A multivariate analysis of variance examined differences between quiz scores as a function of teacher. Since quiz scores were not necessarily comparable across observations, a repeated measures analysis was not used. While the multivariate test for teacher was significant [$F(5,44)=8.86$, $p<.001$] the univariate tests indicated that Fred's class outperformed Mary's for Desert Quiz 2 ($p<.05$), Mary's class outperformed Fred's for Desert Quiz 1 ($p<.01$) and Desert Quiz 4 ($p<.01$). No differences were seen for Desert Quiz 3 or the Rainforests Quiz.

Further analysis revealed that teacher's rank (the original rank ordered list of students used by each teacher to assign students to groups) was a significant predictor of student achievement on all four Desert quizzes ($p<.05$ for both the multivariate and univariate regressions). Student sex was not a significant predictor on any of the five quizzes ($p>.05$).

Overall then, the classes did not appear to differ on their quiz scores nor did boys and girls perform differently on these measures. Teachers did appear to be able to predict student achievement as indicated by the significant relationship of their rankings to students' quiz scores. This findings provided support for the validity of teachers' construction of mixed-ability groups since ranking was the measure used to determine group membership.

Verbal Interaction Results

The coding scheme devised by Cooper, Marquis & Ayers-Lopez (1982) was expanded and modified for application to cooperative learning sessions. The coding scheme is presented in Table 7.

Table 7 About Here

As shown in Table 7, the coding scheme was four levels deep. The first level identified the type of interaction as a Teaching or Learning Move, Response, Pacing or Management Statement of an Evaluation. Additionally student instances of jabbering (more than 30 seconds of off-task interaction) and sentences which were uncodable (Fragment) were noted. The second level provide further detail concerning the specific nature of the communication. For example Teaching Moves could be Attention Focussing, Offers of Assistance, The Asking of Questions, Providing Directives, Making Relevant Comments, etc.

The third level concerned the content of the interaction. Procedural content referred to physical set-up items such as asking for pencil, paper, moving chairs, etc. Factual content concerned specific facts such as the kind of trees found in the desert or the names of animals. Conceptual content included interactions focussed on comparative categories or decision-making criteria, while Process content involved questions and statements about the task and the requirements involved with completing it. Finally Answers refers to interactions conducted once the answer sheets were received at the end of the cooperative learning episode.

The fourth level concerned the specificity and directness of questions and directives. The fourth level data is currently undergoing analysis.

For each of the 36 transcripts, instances of each three level deep communication were computed for each student across the session. These data were then aggregated to created two-level and single level summaries. Multivariate analyses were conducted on the main measures (that is single level Teaching, Learning, Responses, Pacing, Management and Evaluation responses. Second level interactions were analyzed by univariate tests as the number of measures made multivariate tests difficult to interpret.

First Level Analysis for Deserts Sessions

Table 8 presents means and standard deviations for the main single-level measures for both classes across sessions one and two of the Deserts units.

Table 8 About Here

Perusal of the data in Table 8 indicates a number of notable trends. Verbal interaction as identified by the Teaching, Learning and Responses measures shows a drop-off across sessions for Fred's class but not for Mary's. Mary's class shows increases in all but the Pacing and Management measures. The multivariate tests provided statistical support to these observations.

The analysis of these six dependent measures with teacher as the independent variable and sessions as the repeated measure indicated strong significant effects for Teacher [$F(6,40)=3.52$, $p<.01$], Session [$F(6,40)=11.91$] and Teacher by Session interaction [$F(6,40)=4.00$. $p<.01$]. The univariate tests indicated that the Teacher effect was related to Teaching Moves and Responses where Mary's class made significantly more Moves and Responses than Fred's class. Session effects were significant ($p<.05$) for Pacing and Managment. The interaction effect was seen for Learning which indicated a significant decrease for Fred's class as compared to the increase for Mary's class.

Second Level Verbal Interaction Measures

Table 9 provides mean scores for second level measures for students enrolled in Fred's and Mary's class across sessions. Due to the large number of measures, the standard deviations are not presented. Results of the univariate analyses of variance with teacher, sessions and teacher by sessions interaction are presented as well. Since length of sessions in the Desert's unit did not differ significantly between teachers (Fred's sessions lasted, on average 27 minutes; Mary's 29 minutes) the effects of time upon differences were ignored.

Table 9 About Here

Quite simply the data in Table 9 suggest that concerning Teaching Moves the main difference between Fred's and Mary's classes was in the number of contributions that children made to the solution of the task. Mary's children interacted more overall and the task by session interaction indicates that this difference increased over sessions.

Table 10 provides mean data for the Learning Moves second-level variables. These measures related to specific requests for help or information by students.

Table 10 About Here

Students tended to limit their learning moves to specific questions designed to elicit information. On average only about one question per student per session was asked but each student tended to ask about five questions in Fred's class, and about eight in Mary's. The significant difference in Questions between classes indicate that interaction was greater in Mary's class but more importantly interaction maintained itself while Fred's class dropped off greatly during session 2. In contrast nonspecific requests for help were higher in Fred's class for the first session and virtually disappeared in both classes in the second session.

Table 11 presents data related to Student Responses. During the sessions students frequently gave a series of responses to a single question. Only rarely, were non-informative responses given to students' questions.

Table 11 About Here

The results indicate that while the classes did not differ in their number of responses, they did differ significantly in their acknowledgement of statements as well as the number of Argument/Corrective statements. This provides further evidence of increased interaction in Mary's classes.

Pacing, Management and Other Measures

Table 12 presents the remaining main measure collected during the cooperative learning episodes.

Table 12 About Here

These results indicated a fairly consistent pattern: as students became more experienced with the task, the number of Pacing and Management-related statements decreased. Students both praised and criticized the task more in Fred's class but the rate of such responses dropped off during the second session. For the most part, the results indicate that a very small percentage of student interaction was related to non-task specific content.

First Level Verbal Interactions for Rainforests Session

Table 13 presents means and standard deviations for the main single-level measures for the Rain Forests Unit for both classes. Data were collected from each group at only one session. Since length of session did not differ significantly (Fred=37 minutes; Mary=33 minutes) the effects of time upon verbal interaction were not statistically adjusted for length of session.

Table 13 About Here

Data from the Rainforests cooperative learning sessions may be seen as a third observation of the cooperative learning process in the two classrooms. Levels of interaction continued to fall off for Fred's class while Teaching and Response rates continued high in Mary's class. The multivariate tests on the six dependent measures with class as the independent measure confirms these observations. The analysis of these six dependent measures with teacher as the independent variable indicated a strong significant effect for Teacher [$F(6,41)=3.58$, $p<.01$].

The univariate tests indicate that the Teacher effect was significant for Teaching Moves ($p<.001$), Responses ($p<.01$) and Management $p<.01$). The overall pattern is one of increased interaction in Mary's class as compared to Fred's.

Second Level Verbal Interaction Measures for Rainforests

Table 14 provides mean scores for second level measures for students enrolled in Fred's and Mary's class across sessions. Results of the univariate analyses of variance between teachers are presented as well.

Table 14 About Here

The data in Table 14 provide striking evidence of the differences in interaction in the learning groups found in Fred and Mary's classes by the time of the Rainforests unit. Three times as many Teaching Contributions were made in Mary's class, and her students were more likely to focus attention on the task and make relevant comments about the task.

Table 15 provides mean data for the Learning Moves second-level variables. These measures related to specific requests for help or information by students.

Table 15 About Here

Again, students tended to limit their learning moves to specific questions designed to elicit information. While Mary's students again tended to ask about eight questions each during the session as compared to five in Fred's class this difference was not statistically significant. No differences were seen in Learning Moves by students in the two classes.

Table 16 presents data related to Student Responses. During the sessions students frequently gave a series of responses to a single question. Only rarely, were non-informative responses given to students' questions.

Table 16 About Here

Students within classes differed significantly in their acknowledgement of statements as well as the number of Questions asked in response to statements and Argument/Corrective responses. This provides still further evidence of greater interaction in Mary's classes.

Table 17 presents the remaining main measure collected during the cooperative learning episodes.

Table 17 About Here

Mary's class tended to make more Management statements indicating greater interaction. No differences were apparent between classes in Pacing and Evaluation statements. However, on average, each of Fred's students engaged in three episodes of 30 second or more off-task behaviour (Jabber) providing support for the view that the task had become boring for them.

These results provide further support for the view that the high structure of the task in Fred's class inhibited interaction and supressed interest and enthusiasm. The differences in interaction, specifically Teaching Moves and Responses were large and were correlated with achievement differences between Fred and Mary's classes.

Relationship of Verbal Interaction Patterns and Student Achievement

The following sections explore the relationship of the verbal interaction measures to variations in student achievement. First, the overall relationship between number of utterances and student achievement is explored at the student level. Second verbal interactions are correlated with achievement using the student group as the unit of analysis. Finally, the relationships observed for the whole sample are re-examined within each class. Additional analyses examine the differences in achievement and verbal behavior as a function of student ability (as determined by teachers' initial rankings of students) and sex of student.

Student Level Analysis of Relationship Between Verbal and Achievement Behavior

The number of utterances made by students was correlated with their achievement. Since sessions differed in length, partial correlations, controlling for length of sesssion, were also computed.

Forty-nine students contributed data at both sessions. There was moderate consistency in the results. The between session correlation of session length was .36. Between session correlations for verbal interactions were: Teaching Moves: .59; Learning Moves: .22; Responses: .50; Pacing Statements: .54; Management Statements: .27; and Evaluations: .01.

Table 18 presents the correlations of verbal interaction measures from both sessions with student achievement. It should be noted that the verbal measures were obtained from the Deserts sessions since that is where students spent virtually all their learning time. Forty students contributed complete data sets. Since some of the worksheets were shorter than others and since length of session was significantly related to Teaching Moves: .44-Session I/.60 Session II as well as Learning Moves: .58/.46 and Responses: .59/.46, the correlations presented have partialled out the effects of length of session.

Table 18 About Here

Relationships tend to be of low magnitude. Some evidence is seen that Session II measures of Teaching, Learning and Responses were related to posttest Decision and posttest and retention Comparison Making scores. A cluster of positive significant correlations was seen for Deserts Quiz 3 and 4 for these same measures. These findings support the view that students that interacted to a greater extent were more likely to score higher on achievement indices.

A more appropriate analysis used the group as the unit of measurement. Mean scores for both verbal and achievement indices were calculated and correlated. The very small number of groups (n=12) make the attaining of statistical significance difficult, yet very high correlations were seen. Table 19 presents the results of this analysis. Note that the effects of

length of session were controlled by means of partial correlation.

Table 19 About Here

As shown the correlations are strikingly high for the Session II and posttest and retention Decision and Comparison-Making achievement results. These findings offer strong support for the view that groups in which greater interaction took place showed significantly higher achievement on the measures of interest. The correlations of these verbal indices with the quiz measures are puzzling, however; in some cases positive and other negative. Groups in which many Management and Pacing statements were made showed lower achievement.

Stability of Relationships Within Classes

Expanding the results just presented, the consistency of relationship between amount of interaction and student achievement was examined within classes focussing upon the Decision and Comparison Making measures. Table 20 presents these results.

Table 20 About Here

Analysis of these data is very suspect as the number of groups contributing to each correlation was only six. In Fred's class, Teaching Moves were related to retention for both measures and to posttest Decision-Making. In essence those groups interacting more had higher achievement. In Mary's class posttest findings were fairly consistent in that greater Teaching Moves, Responses and Learning Moves were related to higher retention and posttest scores in Comparison-Making and higher retention in decision making. Further analyses will profile the specific characteristics of the higher achieving groups in each classes.

Relationship of Student Ability and Sex to Verbal Behavior

The literature suggests that student ability and/or sex would be related to verbal behaviour. One analysis correlated sex and student ranking (n=49) with these indices. The correlation of rank and verbal behavior was not significant for any of the main measures except for Learning Moves. Student ability correlated negatively with asking questions ($r = -.24$) during the first session of the Deserts unit, particularly specific questions (-.28) and asking for help (-.24). These correlations were exceedingly modest and did not occur for Session II. No other ability relationships were seen.

No relationship was observed between sex and any of the first or second level verbal measures during Session I or II.

An additional analysis examined sex and ability effects. Students were placed into high, moderate and low ability groups based upon a three layer split on teacher rankings. Differences in verbal behavior among students so classified were examined as were the effects of sex by means of analysis of variance on the main verbal indices of Teaching, Learning, Responses, Pacing,

Management and Evaluation. Since students were purposefully stratified into groups upon the basis of ability and sex the effects of length of session were not covaried out. Table 21 summarizes the results of this analysis for Session 1.

Table 21 About Here

The only significant findings was the significant interaction indicating that girls were especially likely to make more Teaching Moves and Responses when they were of high ability and less likely to do so when they were of low ability. These findings were not significant for Session II effects (not tabled).

In summary then, Sex and Ability differences were not salient for the main study measures. Further analysis will examine the second and third level measures to ascertain the presence of these differences among the children of the study.

Social Data

Students were asked at the beginning of their cooperative learning sessions and, again at the end, three questions:

> (a) You have some time to fool around and play games. Of all the people in your class, who would you like to join you? [Who]
>
> (b) You have some school work to do. Your teacher says you can form small groups and work on it together. Of all the people in your class, which ones would you like to have in your group? [Group]
>
> (c) You have some school work to do and you don't know how to do it. Your teacher says you can ask for help from others. Of all the people in your class, which ones would you ask for help? [Help]

At each time of responding students were asked to name six students.

The first analysis examined the relationship between teacher ranking of ability and the number of times each child was chosen by other students at both times of testing. An additional analysis examined whether boys were more likely to be chosen by girls. Students were then placed into high, moderate and low ability groups and analysis of variance examined possible interaction of social status with sex. Finally, crosstabulations were made to see if students that had worked within specific groups were more likely to be chosen at the end of the cooperative learning sessions than was the case prior to the cooperative learning group situation.

Sex was unrelated to the number of times that students were chosen. Rank, however, was significantly related to PreWho, -.25; PreGroup, -.41; PreHelp, -.52; as well as to the post measures: RWho, -.24; RGroup, -.42; and RHelp, -.53.

The analysis of variance of high, moderate and low ability groups by sex revealed that significant main effects were seen for Ability Groups for PreGroup and PreHelp but not PreWho. Sex effects were not observed. Similar findings were seen for the post measures in that students of higher ability were more likely to be chosen for inclusion in learning groups, asking for help, but not to fool around with.

Crosstabulations with Learning Groups

The number of times that each student was chosen foı each measure by members of his or her own group was calculated. The maximum (for a 4 member group) would be 12 (each of 3 students chosen by all 4 members of the group).

Table 22 shows the number of times students within each group were chosen by members of each group.

Table 22 About Here

As shown in Table 22, the results did not support the view that students would be more likely to choose members of their own groups after their work experience than was the case at the beginning. A reason for this may be the very high stability of popularity ratings within the classes. Perusal of rank orderings of student popularity indicated that students ratings were remarkably consistent over time with correlations between the beginning and end of the study of .75 for Who; .69 for Group; and .88 for Help. These results suggest that the Deserts group experience did not lead to significant shifts in popularity among students.

CONCLUSIONS

The results from the two classrooms show that the achievement of students was much higher in the low structured than in the high structured version of cooperative learning. The results also show that achievement is closely linked to the amount and type of verbal interaction within the cooperative groups.

Study 3 contributes a great deal to our undersatnding of the effects of cooperative learning. The findings strongly suggest that it is the verbal interaction that is crucial to the effectiveness of the approach. Previous investigators have attributed the effectiveness of the method to the manipulation of the reward structures. But in Study 3 the reward structures did not vary while achievement did.

The completed analyses indicate partial support for the hypotheses emerging from the literature on naturally occuring groups but there are a number of further lines of investigation to be pursued with the Study 3 data. As noted earlier in the report, the pattern of differences between the classes has yet to be explored in terms of level 4 interaction indicators and the correlations of achievement with verbal interaction have yet to be addressed for levels 2-4. Another promising line of investigation involves searching for specific sequences of interactions and linking these to achievements.

SECTION 3: DISSEMINATION ACTIVITIES

Dissemination of Study 1 Outcomes

Study 1 involved 17 grade 4 classrooms in a large school system. The results of the investigation indicated that the thinking skills treatments led to substantially more learning of problem solving skills than the familiar method in which skills are addressed incidentally while covering the content. Interviews with teachers during and after the study indicated there was strong support for the whole class and cooperative group methods. These findings were disseminated in the following ways:

1. All teachers who were involved in the field test received a short report from the researchers that outlined the main findings from the study and described the practical implications for classrooms. Included in each report were the mean scores on each measure for each treatment. Each teacher also received a confidential appendix that gave the mean scores for his or her own class. Suggestions were made for drawing fair comparisons between these confidential scores from individual classes and the overall results. Teachers were invited to call the investigators (collect) if they wished to discuss the study, and the results from their own class, in more detail.

2. The researchers gave a written and oral presentation of the study and its results to the senior administration of the Board. The administration expressed satisfaction in the findings and encouraged the investigators to undertake the qualitative study described in Section 2 of this report. The administration also suggested that the units developed for the study be revised and be part of the grade 4 Environmental Studies program of the Board.

3. The findings of the study and the practical implications for improving children's learning of intellectual skills were presented to the Curriculum Committee of the Board. The trustees responded enthusiastically to the investigation and encouraged the administration to continue working with the researchers on the qualitative study described in Section 2 of this report.

4. The researchers met with a revision team consisting of some of the teachers that were involved in the production of the original units and several of the field test teachers. After a careful assessment of the student achievement data, the comments of field test teachers and the observations of the researchers during the field test, the two units were revised.

The main changes to the units involved the worksheets for students. The number of practice opportunities was increased but each exercise was made less time consuming. (This was accomplished by dividing the information on the content in the original worksheet into two sets.) In addition occasional minor improvements were made to the worksheet directions to ensure their

clarity and to emphasize to students that following the same process can produce individual differences in the outcome products.

At the time of writing the revision team is developing a broader dissemination strategy for the Board in which inservice will be provided to grade 4 teachers who will be using the units. The initial outline of this plan calls for the division of the grade 4 teachers (n=200+) into four groups. In October, 1987 each group will be given a half day inservice by the principal investigators focussing on the teacher's task in improving children's complex intellectual skills. Each teacher will then teach one unit and incorporate intellectual skill instruction into other areas of the curriculum. In November each group will meet again with the investigators for a further half day inservice to develop additional activities in a "make and take" workshop. Plans for extensions to other grades and subjects are also under development. The implementation of thinking skills will be monitored and stimulated by curriculum improvement teams in each school. These teams, consisting of the principal, teacher-librarian and grade 4 teachers in 1987-88 will be expanded to include grade 5 teachers in 1988-89 and grade 6 teachers in 1989-90.

The revised units are included in Appendix 1 and 2 of this report.

5. The findings from Study 1 (and of Study 2) were reported by the investigators to other researchers attending the Canadian Society for the Study of Education conference in Hamilton in June of 1987. This paper will be revised and submitted to a refereed academic journal.

6. The findings from Study 1 and 2 will be included in an MEd course to be offered in Toronto in the summer of 1987, and in Peterborough in the fall, on programming and implementing complex curriculum objectives.

Dissemination of Study 2 Findings

Study 2 involved 14 grade 4 classrooms in a small school system. The study was intended to be a replication of Study 1 and produced very similar findings: Students given direct instruction in problem solving, either in a whole class or cooperative learning group format, substantially outperformed students given a content driven approach. Teachers were enthused about the direct instruction method and expressed a willingness to continue teaching skills in this manner.

1. All teachers who were involved in the field test received a short report from the researchers that outlined the main findings from the study and described the practical implications for classrooms. Included in each report were the mean scores on each measure for each treatment. Each teacher also received a confidential appendix that gave the mean scores for his or her own class. Suggestions were made for drawing fair comparisons between these confidential scores from individual classes and the overall results. Teachers were invited to call the investigators (collect) if they wished to discuss the study, and the results from their own class, in more detail.

2. A report on the findings was made to the senior administration of the Board. The administration, in consultation with the investigators, decided that the three units developed for the project would be included in the Social Environmental Studies Program of the Junior Division. Since teachers had expressed satisfaction with the materials and concern with their length, it was decided that a) the units would not be revised, b) that the units would be assigned across the Division such that Deserts would be taught in grade 4, Grasslands in grade 5 and Rainforests in grade 6, and c) that teachers would be encouraged to develop additional exercises to reinforce students problem solving skills in the remaining units of each grade. These units are contained in Appendix 3, 4 and 5.

3. A one hour inservice was provided to all Junior Division teachers briefly describing the project, its outcomes and the implications for teachers. Response to this presentation was so favourable that a further follow up session was organized.

4. A full day inservice was held for all Junior Division teachers in the Board. This inservice consisted of input from the investigators on how to improve children's thinking and a series of small group exercises. Each small group was directed by a "veteran" teacher--a member of the field test group or a member of the collaborative writing team that assisted the investigators in the production of the original units. An additional half day inservice was given to the "veterans" to prepare them for their training tasks.

Each teacher attending the session was given a short manual describing how to teach thinking skills. The input portions of the day followed this manual closely and illustrated the main themes of developing ideas about student growth, developing teaching strategies and assessment methods. The small group exercises provided practice on each of these themes. The main small group exercise required each group to produce 4-6 student activities that would reinforce the skills introduced in the three units produced in the project.

The resulting packages of activities were collected from each group. A revision committee, (consisting of the original writing group, a representative of the field test group, an administrator and the investigators) spent a full day revising the exercises. The exercises for the 15 units were then divided into grade packages and distributed to teachers for immediate use. These exercises are contained in Appendix 6.

5. A presentation was made to the principals in the County describing the tasks now being undertaken by their Junior Division teachers, along with specific suggestions about how principals could support and encourage these activities. Special attention was given to the role of the principal in providing or coordinating ongoing motivation, ongoing inservice and ongoing monitoring of progress. Principals were also invited to explore alternative ways of extending direct instruction in thinking skills to other grades and subjects.

6. A presentation was made to all teachers of grade 7 to OAC History courses in the Board. Since the expectations for the recently released Guidelines for History and Contemporary Studies correspond very closely to the

instructional approaches attempted in the Junior Division project, explicit linkages were drawn between Study 2 and the new guidelines.

7. As noted above in the discussion of Study 1, the findings from Study 2 will be disseminated in several ways to academic users: a paper was presented at the CSSE conference; an article will be submitted to a refereed academic journal, and the findings will be incorporated into an M.Ed. course taught in Toronto in the summer and in Peterborough in the fall.

Dissemination of Study 3 Findings

Study 3 involved an intensive investigation of cooperative learning group processes in 2 grade 4 classrooms in the same large school system that participated in Study 1. The complex set of findings are reported in Section 2 of this report.

At the time of writing this report the results of this investigation have been presented, in part, to the two teachers who taught the program--they have also been involved in the interpretation of the audio and video tapes. It is anticipated that the findings from this investigation will be presented to the senior administration of the Board, and that a dissemination strategy will be developed to implement the cooperative learning groups strategy with Junior Division teachers.

References

Bargh, J. & Schul, Y. On the cognitive benefits of teaching. Journal of Educational Psychology, 1980, 72(5), 593-604.

Bloom, B. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 1984, 13(6), 4-16.

Burns, R. & Lash, A. A comparison of activity structures during basic skills and problem-solving instruction in seventh-grade mathematics. American Educational Research Journal, 1986, 23(3), 393-414.

Cohen, P., Kulik, J. & Kulik, C. Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 1982, 19(2), 237-248.

Cooper, C., Marquis, A. & Ayers-Lopez, S. Peer learning in the classroom: Tracing developmental patterns and consequences of children's spontaneous interactions. In L. Wilkinson (Ed.) Communicating in the classroom. New York: Academic, 1982.

Devin-Sheehan, L., Feldman, R. & Allen, V. Research on children tutoring children: A critical review. Review of Educational Research, 1976, 46(3), 355-385.

Eder, D. & Felmlee, D. The development of attention norms in ability groups. In P. Peterson, L. Wilkinsion & M. Hallinan (Eds.) The social context of instruction: Group organization and group process. New York: Academic Press, 1984.

Foster, G. & Penick, J. Creativity in a cooperative group setting. Journal of Research in Science Teaching, 1985, 22(1), 89-98.

Good, T. & Marshall, S. Do students learn more in heterogeneous or homogenous groups? In P. Peterson, L. Wilkinson, & M. Hallinan (Eds.) The social context of instruction: Group organization and group process. New York: Academic Press, 1984.

Johnson, R., Johnson, D. & Stonne, M. Comparison of computer assisted cooperative, competitive, and individualistic learning. American Educational Research Journal, 1986, 22(3), 382-392.

Johnson, D., Maruyama, G. Johnson, R., Nelson, D. & Skon, L. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. Psychological Bulletin, 1981, 89(1), 47-62.

Latane, B., Williams, K. & Hawkins, S. Many hands make light the work: Causes and consequences of social loafing. Journal of Personality and Social Psychology, 1979, 37(6), 822-832.

Lew, M. Mesch, D., Johnson, D. & Johnson, R. Positive interdependence, academic and collaborative-skills group contingencies, and isolated students. American Educational Research Journal, 1986, 23(3), 476-488.

Lindow, J., Wilkinson, L. & Peterson, P. Antecedents and consequences of verbal disagreements during small-group learning. Journal of Educational Psychology, 1985, 77(6), 658-667.

Maynes, F. & Ross, J. Integrating school curricula with social concerns. Alberta Journal of Educational Research, 1984, 30(3), 179-193.

McDonald, B., Dansereau, D., Garland, J., Holley, C. & Collins, K. Pair learning and transfer of text processing skills. Paper presented at the American Educational Research Conference, 1979.

Michaels, J. Classroom reward structures. Review of Educational Research, 1977, 47(1), 87-98.

O'Donnell, et al. Cooperative writing: Direct effects and transfer. Paper presented at the American Educational Research Association Conference, 1985.

Okebukola, P. & Ogunniyi, M. Cooperative, competitive and individualistic science laboratory interactions patterns--effects on students; achievement and acquisition of practical skills. Journal of Research in Science Teaching, 1984, 21(9), 875-884.

Peterson, P. & Fennema, E. Effective teaching, student engagement in classroom activities and sex related differences in mathematics. American Educational Research Journal, 1985, 22(3), 309-355.

Peterson, P., Janicki, T. & Swing, S. Ability X treatment interaction effects on children's learning in large and small group approaches. American Educational Research Journal, 1981, 18(4), 453-473.

Peterson, P. & Swing, S. Student cognitions as mediators of the effectiveness of small group learning. Paper presented at the American Educational Research Association Conference, 1985.

Peterson, P., Wilkinson, L., Spinelli, F. & Swing, S. Merging the process-product and the sociolinguistic paradigms: Research on small group processes. In P. Peterson, L. Wilkinson, & M. Hallinan, (Eds), The social context of instruction: Group organization and group process. New York: Academic Press, 1984.

Rich, Y., Amir, Y., Ben-Ari, R., & Mevarech, Z. Cooperative strategies for mixed-ethnic groups in the Israeli desegregated classroom. Paper presented at the IASCE Conference, Regina, 1985.

Robinson, F., Ross, J. & White, F. Curriculum development for effective instruction. Toronto: OISE Press, 1985.

Ross, J. & Maynes, F. Teaching problem solving. Toronto: OISE Press, 1982.

Ross, J. & Maynes, F. Teaching problem solving: An instructional design strategy. Canadian Journal of Education, 1983, 8(2), 155-173.

Ross, J. & Maynes, F. Retention of problem solving performance in school contexts. Canadian Journal of Education, 1985, 10(4), 383-401.

Sharan, S. Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. Review of Educational Research, 1980, 50(2), 241-272.

Slavin, R. Classroom reward structure: An analytic review. Review of Educational Research, 1977, 47(4), 633-652.

Slavin, R. Cooperative learning. Review of Educational Research, 1980a, 50(2), 315-342.

Slavin, R. When does cooperative learning increase student achievement? Psychological Bulletin, 1983, 94(3), 429-445.

Slavin, R. Using student team learning: The Johns Hopkins team learning project. 2nd edition Baltimore: Johns Hopkins University, 1980b.

Slavin, R. Using student team learning: The Johns Hopkins team learning project. 3rd edition Baltimore: Johns Hopkins University, 1986.

Slavin, R. & Karweit, N. Mastery learning and student teams: A factorial experiment in urban general mathematics classes. American Educational Research Journal, 1984, 21(4), 725-736.

Slavin, R. & Karweit, N. Effects of whole class, ability grouped, and individualized instruction on mathematics achievement. American Educational Research Journal, 1985, 22(3), 351-367.

Stodolsky, S. Frameworks for studying instructional processes in peer work-groups. In P. Peterson, L. Wilkinson, & M. Hallinan (Eds.), The social context of instruction: Group organization and group process. New York: Academic Press, 1984.

Webb, N. A process-outcome analysis of learning in group and individual settings. Educational Psychologist, 1980, 15(2), 69-83.

Webb, N. Student interaction and learning in small groups. Review of Educational Research, 1982, 52(3), 421-445.

Webb, N. Student interaction and learning in small groups: A research summary. In R. Slavin, S. Sharan, L. Kagan, R. Hertz-Lazarowitz, N. Webb & R. Schmuck (Eds.) Learning to cooperate: Cooperating to learn. New York: Plenum Press, 1985.

Webb, N. & Cullian, L. Group interaction and achievement in small groups: Stability over time. American Educational Research Journal, 1983, 20(3), 411-423.

Webb, N. & Kenderski, C. Student interaction and learning in small group and whole class settings. In P. Peterson, L. Wilkinson & M. Hallinan (Eds.) The social context of instruction: Group organization and group process. New York: Academic Press, 1984.

Footnotes

1. The procedure was developed by Dennis Raphael in consultation with Tom Moore of the OISE Instructional Resources Development Unit, Shawn Moore of the Computer Literacy Project and Bill McKellin of the Clasroom Interaction Project. Helpful suggestions were also made by Durham Region Board of Education officials. A review of the research literature also surfaced several procedural suggestions.

TABLE 1

UNADJUSTED MEANS AND STANDARD DEVIATIONS FOR STUDY 1
CLASS AS UNIT OF ANALYSIS

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
SKILL				
Comparative Thinking: Matching Information	pre	1.46 (.14)	1.64 (.26)	1.02 (.55)
	post	1.59 (.25)	1.83 (.19)	1.67 (.34)
	retention	1.69 (.23)	1.89 (.16)	1.82 (.17)
Comparative Thinking: Use of Categories	pre	0.01 (.02)	0.01 (.03)	0.12 (.19)
	post	0.03 (.05)	1.25 (.82)	1.06 (.69)
	retention	0.04 (.10)	1.21 (.74)	1.09 (.84)
Comparative Thinking: Interpreting a Comparison	pre	0.39 (.25)	0.33 (.33)	0.25 (.19)
	post	0.35 (.23)	0.51 (.40)	0.35 (.29)
	retention	0.52 (.29)	0.53 (.28)	0.45 (.40)
Decision-Making: Generating Alternatives	pre	0.33 (.10)	0.24 (.12)	0.30 (.19)
	post	0.39 (.21)	0.72 (.31)	0.66 (.24)
	retention	0.33 (.16)	0.80 (.23)	0.67 (.26)
Decision-Making: Developing Criteria	pre	0.09 (.13)	0.14 (.18)	0.17 (.28)
	post	0.04 (.06)	2.76 (1.90)	2.04 (1.74)
	retention	0.11 (.10)	3.01 (1.66)	1.93 (1.79)
Decision-Making: Summarizing Information	pre	0.00 (.00)	0.02 (.05)	0.04 (.07)
	post	0.00 (.00)	0.44 (.42)	0.42 (.37)
	retention	0.06 (.06)	0.48 (.34)	0.32 (.31)
AFFECT				
Love of Learning	pre	3.61 (.19)	3.42 (.25)	3.47 (.38)
	retention	3.47 (.28)	3.44 (.30)	3.45 (.33)
Perceived Competence	pre	62.58 (6.86)	62.38 (8.58)	62.57 (5.80)
	retention	62.75 (7.23)	68.02 (6.58)	69.61 (7.16)
Rainforests	pre	0.54 (.09)	0.55 (.12)	0.52 (.11)
	post	0.65 (.06)	0.67 (.10)	0.68 (.04)
	retention	0.63 (.07)	0.68 (.09)	0.64 (.07)
Deserts	pre	0.57 (.14)	0.50 (.08)	0.56 (.10)
	post	0.77 (.09)	0.78 (.09)	0.73 (.15)
	retention	0.75 (.09)	0.73 (.11)	0.73 (.09)

TABLE 2

UNADJUSTED MEANS AND STANDARD DEVIATIONS FOR STUDY 1
STUDENT AS UNIT OF ANALYSIS

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
AFFECT				
Love of Learning	pre	3.58 (.53)	3.40 (.59)	3.41 (.67)
	retention	3.53 (.61)	3.41 (.67)	3.47 (.68)
Perceived Competence	pre	65.09 (17.25)	65.76 (15.67)	67.25 (13.86)
	retention	68.15 (17.01)	70.32 (15.70)	71.18 (15.52)
KNOWLEDGE				
Rainfall	pre	.53 (.23)	.55 (.23)	.53 (.25)
	post	.65 (.20)	.69 (.23)	.68 (.21)
	retention	.61 (.20)	.69 (.22)	.65 (.24)
Deserts	pre	.55 (.27)	.51 (.25)	.57 (.23)
	post	.76 (.21)	.77 (.18)	.76 (.22)
	retention	.73 (.22)	.74 (.22)	.75 (.21)
SKILL				
Comparative Thinking: Matching Information	pre	1.42 (.79)	1.64 (.64)	1.08 (.89)
	post	1.54 (.72)	1.83 (.47)	1.68 (.68)
	retention	1.69 (.58)	1.85 (.49)	1.82 (.48)
Comparative Thinking: Use of Categories	pre	.01 (.10)	.01 (.12)	.08 (.37)
	post	.04 (.25)	1.25 (.95)	1.07 (.98)
	retention	.03 (.22)	1.20 (.97)	1.10 (.99)
Comparative Thinking: Interpreting a Comparison	pre	.39 (.49)	.30 (.46)	.25 (.43)
	post	.36 (.48)	.50 (.50)	.29 (.46)
	retention	.58 (.50)	.52 (.50)	.43 (.49)
Decision-Making: Generating Alternatives	pre	.33 (.47)	.26 (.44)	.33 (.47)
	post	.40 (.49)	.73 (.45)	.69 (.49)
	retention	.35 (.48)	.81 (.39)	.65 (.48)
Decision-Making: Developing Criteria	pre	.09 (.46)	.16 (.56)	.25 (.49)
	post	.04 (.20)	2.78 (2.29)	2.10 (2.18)
	retention	.11 (.32)	3.06 (2.08	1.84 (2.19)
Decision-Making: Summarizing Information	pre	.00 (.00)	.03 (.16)	.04 (.20)
	post	.00 (.00)	.44 (.50)	.43 (.50)
	retention	.06 (.24)	.51 (.50)	.50 (.46)

TABLE 3

UNADJUSTED MEANS AND STANDARD DEVIATIONS FOR STUDY 2
CLASS AS UNIT OF ANALYSIS

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
SKILL				
Comparative Thinking: Matching information	pre	0.99 (.51)	0.93 (.47)	0.27 (.24)
	post	1.02 (.36)	1.45 (.36)	1.07 (.39)
	retention	1.08 (.41)	1.45 (.45)	1.06 (.53)
Comparative Thinking: Use of Categories	pre	0.09 (.15)	1.08 (.17)	0.05 (.08)
	post	0.07 (.11)	1.25 (.82)	0.99 (.67)
	retention	0.02 (.04)	1.66 (.28)	0.95 (.68)
Comparative Thinking: Interpreting a Comparison	pre	0.29 (.24)	0.22 (.24)	0.08 (.11)
	post	0.25 (.30)	0.44 (.44)	0.63 (.19)
	retention	0.23 (.19)	0.26 (.24)	0.57 (.29)
Experimental Thinking: Identifying the Cause	pre	1.05 (.44)	0.85 (.56)	0.46 (.37)
	post	1.34 (.46)	1.61 (.33)	1.08 (.45)
	retention	1.47 (.25)	1.55 (.27)	1.06 (.49)
Experimental Thinking: Identifying the Effect	pre	0.15 (.15)	0.11 (.10)	0.02 (.04)
	post	0.14 (.18)	0.18 (.11)	0.15 (.17)
	retention	0.18 (.21)	0.32 (.15)	0.22 (.14)
Experimental Thinking: Controlling Variables	pre	0.09 (.13)	0.05 (.08)	0.09 (.13)
	post	0.09 (.09)	2.34 (.56)	1.49 (.81)
	retention	0.13 (.14)	2.13 (.33)	1.34 (.82)
Comparative Thinking: Focussing a Comparison	pre	0.55 (.21)	0.50 (.11)	0.48 (.15)
	post	0.64 (.17)	0.85 (.07)	0.73 (.16)
	retention	0.57 (.19)	0.78 (.06)	0.77 (.17)
Comparative Thinking: Writing a Comparative Paragraph	pre	0.39 (.09)	0.26 (.20)	0.37 (.19)
	post	0.39 (.10)	0.93 (.13)	0.91 (.17)
	retention	0.53 (.12)	0.88 (.24)	0.69 (.15)
Experimental Thinking: Focussing an Experiment	pre	0.24 (.08)	0.14 (.10)	0.18 (.09)
	post	0.30 (.12)	0.70 (.09)	0.29 (.18)
	retention	0.31 (.13)	0.58 (.10)	0.32 (.13)

TABLE 3 cont'd.

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
AFFECT				
Love of Learning	pre	3.73 (.14)	3.50 (.24)	3.65 (.18)
	retention	3.70 (.19)	3.50 (.20)	3.57 (.27)
Perceived Competence	pre	67.66 (6.40)	60.00 (8.17)	61.50 (6.21)
	retention	67.74 (9.15)	65.21 (5.32)	67.10 (6.72)
KNOWLEDGE				
Rainforests	pre	0.64 (.07)	0.66 (.06)	0.62 (.06)
	post	0.74 (.07)	0.78 (.04)	0.72 (.07)
	retention	0.70 (.10)	0.74 (.07)	0.75 (.08)
Grasslands	pre	0.69 (.08)	0.68 (.04)	0.71 (.11)
	post	0.76 (.09)	0.85 (.06)	0.82 (.09)
	retention	0.71 (.06)	0.77 (.08)	0.82 (.07)
Deserts	pre	0.50 (.10)	0.49 (.05)	0.47 (.06)
	post	0.79 (.06)	0.84 (.06)	0.77 (.10)
	retention	0.72 (.06)	0.75 (.09)	0.78 (.09)

TABLE 4

UNADJUSTED MEANS AND STANDARD DEVIATION FOR STUDY 2
STUDENT AS UNIT OF ANALYSIS

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
AFFECT				
Love of Learning	pre	3.74 (.54)	3.50 (.67)	3.60 (.58)
	retention	3.76 (.65)	3.45 (.66)	3.55 (.69)
Perceived Competence	pre	67.43 (14.34)	60.23 (18.15)	61.93 (18.31)
	retention	71.76 (15.90)	65.92 (17.65)	67.99 (19.45)
KNOWLEDGE				
Rainfall Unit	pre	.61 (.21)	.65 (.22)	.62 (.22)
	post	.76 (.19)	.77 (.18)	.73 (.20)
	retention	.72 (.21)	.76 (.20)	.74 (.22)
Grasslands Unit	pre	.69 (.23)	.69 (.21)	.68 (.25)
	post	.74 (.16)	.86 (.16)	.80 (.20)
	retention	.70 (.21)	.78 (.20)	.82 (.19)
Desert Unit	pre	.54 (.19)	.50 (.19)	.48 (.24)
	post	.80 (.17)	.85 (.19)	.76 (.22)
	retention	.73 (.18)	.77 (.21)	.76 (.21)
SKILL				
Comparative Thinking: Matching Information	pre	1.14 (.86)	1.01 (.92)	.26 (.53)
	post	1.18 (.81)	1.41 (.73)	1.16 (.88)
	retention	1.23 (.71)	1.42 (.83)	1.17 (.83)
Comparative Thinking: Use of Categories	pre	.11 (.43)	.04 (.30)	.05 (.22)
	post	.07 (.30)	1.20 (.89)	1.13 (.91)
	retention	.03 (.23)	1.60 (.70)	1.08 (.87)
Comparative Thinking: Interpreting a Comparison	pre	.27 (.45)	.27 (.44)	.07 (.26)
	post	.22 (.42)	.47 (.50)	.64 (.48)
	retention	.22 (.41)	.28 (.45)	.64 (.48)
Experimental Thinking: Identifying the Cause	pre	1.08 (.61)	1.00 (.88)	.54 (.76)
	post	1.30 (.64)	1.65 (.64)	1.11 (.78)
	retention	1.43 (.60)	1.54 (.66)	1.12 (.81)
Experimental Thinking: Identifying the Effect	pre	.19 (.39)	.11 (.32)	.02 (.15)
	post	.19 (.39)	.19 (.40)	.19 (.40)
	retention	.19 (.39)	.33 (.47)	.22 (.42)
Experimental Thinking: Controlling Variables	pre	.12 (.33)	.06 (.31)	.06 (.29)
	post	.09 (.29)	2.37 (1.04)	1.56 (1.25)
	retention	.15 (.49)	2.13 (1.22)	1.44 (1.26)

TABLE 4 continued

VARIABLE	OCCASION	CONTROL	WHOLE	CO-OP
SKILL continued				
Comparative Thinking: Asking Questions	pre	.50 (.35)	.47 (.32)	.45 (.36)
	post	.59 (.32)	.84 (.22)	.73 (.28)
	retention	.60 (.35)	.79 (.24)	.78 (.30)
Comparative Thinking: Writing a Comparative Paragraph	pre	.43 (.35)	.29 (.32)	.34 (.31)
	post	.43 (.29)	.92 (.46)	.92 (.48)
	retention	.57 (.24)	.87 (.52)	.65 (.55)
Experimental Thinking: Asking Questions	pre	.24 (.25)	.15 (.23)	.18 (.23)
	post	.33 (.26)	.70 (.30)	.33 (.37)
	retention	.34 (.27)	.59 (.37)	.33 (.36)

Table 5

Skill Scores for Fred's and Mary's Students

Measure		Fred (n=25) pre	post	retent	Mary (n=22) pre	post	retent
Comparison Making	M	3.84	2.32	2.92	1.86	4.33	4.43
	SD	1.28	1.82	1.87	1.15	.73	.75
Decision Making	M	1.00	1.08	1.64	1.27	4.00	3.86
	SD	.91	1.35	1.65	1.03	.31	.83

Table 6

Quiz Scores for Fred's and Mary's Students

	Fred (n=25)		Mary (n=25)	
Unit	M	SD	M	SD
Deserts 1	22.64	3.36	25.52	3.84
Deserts 2	27.56	2.58	25.60	3.20
Deserts 3	22.92	4.11	22.32	4.42
Desert 4	19.96	3.54	23.16	4.27
Rain-Forests 1	25.64	2.32	25.12	3.92

TABLE 7

CODING SCHEME FOR QUALITATIVE STUDY

The coding takes place on four levels. The first identifies the nature of the communication. The second provides details of the particular communication. The third coding identified the substance or level of the communication. The fourth coding defines the specificity and directness of any directives or questions. The categories marked with an asterisk are from Cooper, Marquis and Ayers-Lopez (1982).

CODING	NATURE OF COMMUNICATION		NOTES
First Coding	Teaching Move	(T)	*ANY spontaneous non-request
	Learning Move	(L)	*ANY request for assistance
	Response	(R)	ANY response
	Pacing	(P)	*E.g., hurry up
	Management	(M)	*student tries to control others
	Evaluation	(E)	ANY positive or negative appraisal
	Fragment	(F)	Uncodeable fragment
	Jabber	(J)	30 seconds of off-task nonsense
	Reading	(B)	Student reads aloud
Second Coding (Moves Only)	Attention Focusing	(1)	*Includes OK, let's go guys, etc.
	Offers of Assistance	(2)	
	Contribution	(3)	E.g., provides a fact required by the exercise
	Directives or Requests for Action	(4)	*E.g., put the temperature here
	Questions or Requests for Information	(5)	*E.g., what's the temperature?
	Requests for Help	(6)	*No specific requests for information
	Relevant Comments	(7)	*On task descriptions not coded as 1-6
	Irrelevant Comments	(8)	*Off task remarks
	Request of Teacher	(9)	Any request of teacher
Second Coding (Responses Only)	Informative Responses	(1)	*Complete specific response
	Uninformative Responses	(2)	*Incomplete or nonspecific answer
	Acknowledgements with Content	(3)	E.g., student repeats "weather"
	Acknowledgement without Content	(4)	*E.g., yeah, ok, etc.
	Question or Directive	(5)	*Answers with a question
	Arguments/Correctives	(6)	*Gives counter-assertion
	No response coding will be obtained through analysis		
Second Coding (Evaluation)	Praising Another	(1)	*E.g., you've got it
	Criticizing Another	(2)	*E.g., that's not right
	Praising Self	(3)	*E.g., I'm good at this
	Criticizing Self	(4)	*E.g., I blew it
	Praising Task	(5)	*E.g., this is neat
	Criticizing Task	(6)	*E.g., this is dumb
Second Coding (Pacing)	Response	(1)	
	Directive	(4)	E.g., "hurry up"
	Query	(5)	E.g., "are people done"
Third Coding (Moves and Responses)	Procedures	(A)	Pencils, lines, seating, etc.
	Factual	(B)	Content: hot, snakes, wet, etc.
	Conceptual	(C)	Categories, Advantages, etc.
	Process	(D)	Task-related questions, etc.
	Answers	(E)	To specific worksheet items
Fourth Coding (Directives and Questions)	Direct, Specific	(1)	
	Direct, Unspecific	(2)	
	Indirect, Specific	(3)	
	Indirect, Unspecific	(4)	

Table 8

Verbal Interaction Scores for Desert Unit Learning Groups

		Fred (n=24)		Mary (n=23)	
Measure		Session 1	Session 2	Session 1	Session 2
Teaching	M	36.32	25.58	42.26	47.35
	SD	21.2	19.81	24.43	21.98
Learning	M	11.04	3.62	8.04	10.13
	SD	8.91	4.32	5.03	7.86
Responses	M	32.17	23.37	33.43	38.96
	SD	18.17	15.43	17.74	19.63
Pacing	M	6.04	2.54	5.48	2.22
	SD	5.47	2.21	4.33	2.19
Management	M	7.33	4.21	6.04	2.91
	SD	3.50	2.55	6.22	3.23
Evaluation	M	2.08	.75	2.00	2.09
	SD	2.96	1.33	3.96	2.29

Table 9

Mean Verbal Interaction Teaching Scores for Desert Unit Groups

	Fred (n=24)		Mary (n=23)		Significance Level of Effect		
	Session		Session				
Teaching	1	2	1	2	T	S	TxS
Attention Focussing	2.60	2.16	2.08	2.58	ns	ns	ns
Assistance Offering	.52	.16	.67	.46	ns	ns	ns
Teaching Contribution	10.92	8.28	15.63	22.67	***	ns	*
Teaching Directive	3.84	3.08	4.38	3.79	ns	ns	ns
Teaching Questions	1.48	1.36	3.08	1.50	ns	ns	ns
Relevant Comment	10.08	6.44	9.71	11.75	ns	ns	ns
Irrelevant Comment	6.84	3.88	6.04	3.46	ns	*	ns

* $p<.05$ *** $p<.001$

Table 10

Mean Verbal Interaction Learning Scores for Desert Unit Groups

	Fred (n=24)		Mary (n=23)		Significance Level of Effect		
	Session		Session				
Teaching	1	2	1	2	T	S	TxS
Request or Directive	.44	.32	.63	.54	ns	ns	ns
Question	8.60	3.24	7.04	9.58	*	ns	***
Help Request	1.40	0.00	.33	.08	*	***	**
Request of Teacher	0.16	.04	.08	.17	ns	ns	ns

* p<.05 ** p<.01 *** p<.001

Table 11

Mean Verbal Interaction Response Scores for Desert Unit Groups

	Fred (n=24)		Mary (n=23)		Significance Level of Effect		
	Session		Session				
Teaching	1	2	1	2	T	S	TxS
Informative Response	12.52	10.16	10.33	10.54	ns	ns	ns
Uninformative Response	1.08	1.12	.83	.75	ns	ns	ns
Acknowledgement (with content)	4.52	2.04	5.46	10.54	***	ns	***
Acknowledgement (without content)	4.52	3.32	5.96	5.00	**	ns	ns
Question	4.08	2.68	3.29	2.92	ns	ns	ns
Argument/ Corrective	5.00	4.40	7.13	8.96	*	ns	ns

* $p<.05$ *** $p<.001$

Table 12

Mean Verbal Interaction Additional Scores for Desert Unit Groups

	Fred (n=24)		Mary (n=23)		Significance Level of Effect		
	Session		Session				
Teaching	1	2	1	2	T	S	TxS
Pacing	5.84	2.52	5.33	2.13	ns	***	ns
Response	1.40	.36	1.08	.25	ns	***	ns
Statement	2.76	1.68	2.73	1.29	ns	***	ns
Directive	.12	.00	.25	.13	ns	ns	ns
Question	1.56	.48	1.38	.46	ns	***	ns
Management	7.60	4.04	6.04	3.04	ns	***	ns
Directive	6.60	3.16	5.42	2.63	ns	***	ns
Statement	.60	.24	.17	.25	ns	***	ns
Evaluation	2.32	.72	1.92	2.00	ns	***	ns
Praising Other	.28	.08	.04	.25	ns	ns	*
Criticizing Other	.80	.48	1.21	1.25	ns	ns	ns
Praising Self	.28	.00	.29	.21	ns	ns	ns
Criticizing Self	.16	.08	.21	.13	ns	ns	ns
Praising Task	.60	.08	.13	.00	*	*	ns
Criticizing Task	.20	.00	.04	.17	*	ns	*
Jabber	2.40	.60	.50	.50	*	ns	ns
Fragments	.40	.24	.96	.04	ns	***	**

* p<.05 ** p< .01 *** p<.001

Table 13

Verbal Interaction Scores for Rain Forests Unit Learning Groups

Measure		Fred (n=25)	Mary (n=23)
Teaching	M	23.60	42.39
	SD	13.22	19.64
Learning	M	6.36	8.13
	SD	6.55	5.44
Responses	M	21.76	33.39
	SD	11.87	16.94
Pacing	M	2.40	3.91
	SD	1.98	4.81
Management	M	2.56	5.39
	SD	2.70	4.27
Evaluation	M	1.12	1.13
	SD	1.72	1.22

Table 14

Mean Verbal Interaction Teaching Scores for Rain Forests Unit Groups

Teaching	Fred (n=24)	Mary (n=23)	Significance Level of Teacher Effect
Attention Focussing	1.04	2.78	**
Assistance Offering	.44	.30	ns
Teaching Contribution	7.52	23.00	***
Teaching Directive	2.28	3.17	ns
Teaching Questions	1.00	.39	ns
Relevant Comment	5.44	8.43	*
Irrelevant Comment	5.88	4.30	ns

* p<.05 ** p<.01 *** p<.001

Table 15

Mean Verbal Interaction Learning Scores for Desert Unit Groups

Learning	Fred (n=24)	Mary (n=23)	Significance Level of Teacher Effect
Request or Directive	.08	.26	ns
Question	5.04	7.57	ns
Help Request	.80	.04	ns
Request of Teacher	0.44	.26	ns

* p<.05 ** p<.01 *** p<.001

Table 16

Mean Verbal Interaction Response Scores for Rain Forests Unit Groups

Responses	Fred (n=24)	Mary (n=23)	Significance Level of Teacher Effect
Informative Response	11.44	8.04	ns
Uninformative Response	.44	.57	ns
Acknovledgement (with content)	4.16	11.00	***
Acknovledgement (without content)	.92	4.52	***
Question	1.20	2.43	*
Argument/ Corrective	3.56	6.83	**

* $p<.05$ *** $p<.001$

Table 17

Mean Verbal Interaction Additional Scores for Rain Forests Unit Groups

Measure	Fred (n=24)	Mary (n=23)	Significance Level of Teacher Effect
Pacing	2.40	3.91	ns
Response	.04	.83	**
Statement	1.88	1.74	ns
Directive	.08	.22	ns
Question	.40	1.13	ns
Management	2.56	5.39	*
Directive	2.20	4.35	*
Statement	.04	.83	***
Evaluation	1.12	1.13	ns
Praising Other	.00	.04	ns
Criticizing Other	.68	.74	ns
Praising Self	.08	.09	ns
Criticizing Self	.08	.22	ns
Praising Task	.20	.00	ns
Criticizing Task	.08	.04	ns
Jabber	3.08	.65	***
Fragments	.36	.35	ns

* p<.05 ** p<.01 *** p<.001

Table 18

Correlations of Verbal Behaviour Indices with Student Achievement
(controlling for length of session)
(n=49)

	Verbal Behavior											
Achievement Measure	Session I						Session II					
	T	L	R	P	M	E	T	L	R	P	M	E
Decision-Making												
Pretest	06	-18	-07	-14	-01	05	00	20	09	02	21	10
Post-test	22	-12	14	09	-21	09	46	38	36	-09	-14	09
Retention	06	-22	-05	-11	-34	-13	18	12	05	-25	-36	01
Comparison-Making												
Pretest	-06	07	-08	13	25	11	-01	-02	05	22	39	00
Post-test	01	-04	00	-07	-34	-01	24	44	19	-19	09	22
Retention	21	-07	-13	-06	-12	22	37	45	42	-10	-01	05
Quiz Scores												
Deserts 1	23	-20	03	-17	-36	18	27	14	13	-20	04	12
Deserts 2	07	-18	-04	-08	33	18	01	-08	08	04	32	-04
Deserts 3	38	-24	11	18	38	27	31	29	35	33	37	13
Deserts 4	27	-27	14	05	-11	13	21	20	20	-07	27	09

with df=41 .25=p<.05 .35 p<.01; .46=p<.001;

Table 19

Correlations of Verbal Behaviour Indices with Student Achievement Group as Unit of Measurement (controlling for Length of Session(n=12)

Achievement Measure	Verbal Behavior											
	Session I						Session II					
	T	L	R	P	M	E	T	L	R	P	M	E
Decision-Making												
Pretest	-02	-69	-26	-25	-05	00	08	01	02	06	04	13
Post-test	41	-13	27	14	-39	05	86	63	53	-34	-26	21
Retention	14	-13	07	-08	-51	-16	78	46	40	-43	-56	32
Comparison-Making												
Pretest	04	03	02	-21	26	34	-65	-30	-35	28	58	-37
Post-test	50	-32	30	-04	-46	13	68	68	66	-43	-03	23
Retention	32	07	38	16	-25	23	84	64	63	-26	-33	18
Quiz Scores												
Deserts 1	-06	-09	02	-10	-22	-21	65	50	26	-54	-44	51
Deserts 2	-36	-08	-29	-04	57	-33	-71	-72	-84	28	09	12
Deserts 3	-31	-34	-38	03	64	-18	-17	-12	-38	58	30	28
Deserts 4	26	-18	07	21	01	-12	34	29	07	-14	13	09

Table 20

Correlations of Verbal Behaviour Indices with Student Achievement Within Fred and Mary's Classes (controlling for Length of Session n=6 for each correlation)

Achievement Measure	Session II Verbal Behavior											
	Fred's Class						Mary's Class					
	T	L	R	P	M	E	T	L	R	P	M	E
Decision-Making												
Pretest	-38	39	12	37	22	35	-04	-69	-28	79	-00	-41
Post-test	84	-02	-13	23	-05	-54	36	73	61	-53	87	-14
Retention	63	-41	-23	-23	-73	-32	10	-68	-40	63	-99	45
Comparison-Making												
Pretest	-08	-17	-73	20	44	59	-18	84	49	-89	89	-41
Post-test	-36	-00	-56	22	89	60	53	36	89	-22	60	-47
Retention	65	32	72	16	-69	-86	49	28	24	00	45	19

Table 21

Summary of Ability and Gender Differences in Verbal Behavior for Session 1 of Deserts

	Student Ability as Rated by Teacher						Significance of Effect		
Ability	Low		Moderate		High				
Measure	Boys	Girls	Boys	Girls	Boys	Girls	A	G	AxG
Teaching	46.6	20.5	31.7	44.1	33.3	50.3	ns	ns	*
Learning	10.8	9.2	9.1	10.1	9.0	6.6	ns	ns	ns
Responses	38.1	16.5	27.3	38.0	31.0	36.2	ns	ns	*
Pacing	6.36	6.5	4.9	5.9	5.7	4.5	ns	ns	ns
Management	6.6	6.7	5.5	5.2	10.8	8.5	ns	ns	ns
Evaluation	1.8	2.0	1.2	2.8	2.2	3.7	ns	ns	ns

* $p<.05$

Table 22

Number of Within-Group Choices Made by Group Members at Pre and Post-Data Collections

	Who			Group			Help		
Group	Pre	Post	Change	Pre	Post	Change	Pre	Post	Change
01	2	0	-	2	3	+	3	3	
02	3	4	+	2	1	-	4	4	
03	4	4		4	2	-	2	1	-
04	4	4		8	7	-	3	9	+
05	6	6		3	3		2	3	+
06	4	4		1	2	+	2	1	-
07	5	3	-	4	3	-	4	3	-
08	2	1	-	1	1		2	1	-
09	3	3		3	2	-	1	2	+
10	2	2		2	4	+	2	2	
11	3	3		4	3	-	3	4	+
12	2	1	-	1	2	+	3	3	

Level	Level Characteristics	Example (Where should the shopping mall be built?)

Level 1 — one course of action/ feeble justification given for it

Build Mall Downtown
near our school so I can go to MacDonald's

Level 2 — several alternative courses of action/ different advantages identified for each alternative

Build the Mall Downtown	Build the Mall East of City
-lots of people already go downtown -good roads -downtown merchants want it here	-good flat land -no buildings would have to be torn down -people in east end would have have a place to shop

Level 3 — several alternative courses of action/different advantages and disadvantages identified for each alternative

	Build the Mall Downtown	Build the Mall East of City
Advantages	-lots of people available -good roads -downtown merchants want it here	-good flat land -no buildings would have to be torn down -would give people in east end a place to shop
Disadvantages	-would have to move some stores -land is very expensive -requires bylaw change	-would have to build new roads -farmland would be lost

Level 4 — several alternative courses of action/explicit criteria systematically applied to each

	Build the Mall Downtown	Build the Mall East of City	Build the Mall West of City
What downtown merchants want	+	-	-
Cost of land	-	+	+
Availability of shoppers	+	+	-
Suitability of roads	+	-	-

Figure 1: Growth Scheme for the Skill of Establishing a Framework for a Decision-Making Problem (from Maynes, F. J., & Ross, J. A., 1984)

WORKSHEET O

Suppose you were stuck in the desert. Your jeep will not go. You have no camel to ride. You must walk to the next oasis. It will take you a whole day.

If you could choose 3 things to have with you. which 3 things on the following list would you choose?

1. one litre of water
2. some dates
3. a knife
4. a sheet of plastic
5. a sleeping bag
6. a package of biscuits
7. a tent
8. a coat
9. a hat
10. sunglasses

1. Make a chart on the attached Worksheet Blank to show how you would decide.
2. Which 3 things would you choose? Why?

Figure 2: Example of a Worksheet

WORKSHEET 0 ANSWER

Compare the charts made by members of the group with the chart below. Are they the same? Why not? Ask about each group member's chart:

1. Are all the possible choices listed entered on the chart? Is there a heading, "Possible Choices" above the columns?
2. Are there 2 rows labelled "advantages" and "disadvantages"?
3. Are there advantages and disadvantages listed for each of the possible choices? (Sometimes it is hard to think of an advantage or a disadvantage for a certain thing.)
4. Have choices been made? Does the reason for the choices make sense given the advantages and disadvantages listed?

	POSSIBLE CHOICES									
	1 litre water	some dates	a knife	a sheet of plastic	sleeping bag	package of biscuits	tent	coat	hat	sun glasses
Advantages	need to drink to survive	good food to keep up energy easy to carry	good if a wild animal threatens easy to carry	could keep you and your things dry if it rains	could keep you warm at night	could keep up energy not hard to carry	could protect against sun or sand if you stop	would help to keep warm at night	would keep the sun off head and face very light to carry	protects eyes from sun
Disadvantages	---	might not like their taste, not essential	might cut yourself	a little heavy to carry	heavy to carry	---	heavy to carry, probably won't stop	a little heavy to carry would not wear during day	---	---

Choice: Biscuits and hat because they have more very good advantages. They give energy and protection. And unlike most of the others, there are no disadvantages. The third one is water. Sunglasses have one advantage and no disadvantages, like water. But water's advantage is much more important.

Figure 3: Example of a Worksheet Answer

	Score
Comparative Problems	
1. Matching information using implicit categories	
.information about each object of comparison will be developed and some of it will be matched	1 mark
.all of it will be matched	2 marks
2. Use of explicit categories to organize information	
.one explicit category used	1 mark
.more than one explicit category	2 marks
3. Interpreting comparisons	
.each comparison summarized as similar or different	1 mark
Decision Making Problems	
1. Identification of alternatives	
.more than one explicit alternative	1 mark
2. Development of criteria	
.primitive strategy for deciding, e.g., voting	1 mark
.identification of explicit advantages	2 marks
.explicit advantages and disadvantages	3 marks
.one explicit criterion applied to all alternatives	4 marks
.more than one explicit criterion	5 marks
3. Summarizing information about values of alternatives	
.use of symbol system, e.g., +'s and -'s	1 mark

Figure 4: Indicators of Problem Solving Skill Used in Study 1

	Score
(A) COMPARATIVE PROBLEMS (as in Study 1)	
(B) EXPERIMENTAL PROBLEMS	
1. Identification of an explicit cause	
.implicit cause--"something" will be changed	1 mark
.explicit cause in specific terms	2 marks
2. Identification of explicit effect	
.effect stated in terms which are more precise than those given in the stimulus	1 mark
3. Controlling variables	
.no explicit variables--"something" will be kept the same	1 mark
.one explicit variable controlled	2 marks
.more than one explicit variable controlled	
(C) FOCUSSING EXPERIMENTAL PROBLEMS	
.cause or effect identified	1 mark
.cause and effect identified	2 marks
.cause, effect and controls identified	3 marks
(D) FOCUSSING COMPARATIVE PROBLEMS	
.two or more things identified in focus	1 mark
.two or more things to be explicitly compared	2 marks
.identifies objects of comparison and one or more categories in which they will be compared	3 marks
(E) WRITING COMPARATIVE PARAGRAPHS [Score divided by 5 to give a range of 0 to 1.0]	
1. Opening sentence	
.identifies what will be compared	1 mark
.identifies objects and categories of comparison	2 marks
2. Use of comparative sentences	
.contains at least one literate sentence with information comparing two things	1 mark
.contains more than one literate comparative sentence	2 marks
3. Closing sentence	
.summarizing statement giving similarities and differences	1 mark

Figure 5: Indicators of Problem Solving Skills Used in Study 2